[54]发明名称 选择B帧编码模式的视频编码方法

[57]摘要

　　H.263十建议中的方案之一,是在改善的PB帧模式里可对一个B帧的宏块按照前向、反向或双向预测模式进行编码,本发明涉及一种对图像序列编码的方法,在编码B宏块的三种可能的预测模式中可确定一个策略来选择出一种预测模式。该策略基于SAD(绝对差值总和)计算和运动矢量相关性,并当发生镜头切换时容许应用反向预测。这里的计算针对原图像实施,可以使计算量较少,并减轻CPU的负担。本发明还涉及一种编码系统,用于实现所述方法,并包括有一种用以存储指令的计算机可读介质,可由指令来实施本方法。
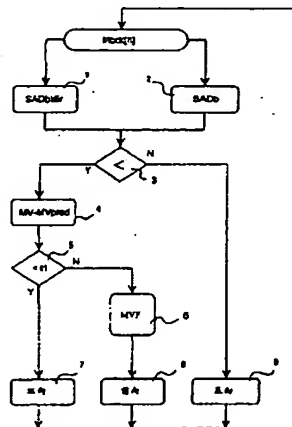
权 利 要 求 书

1. 一种对信源图像序列进行编码的方法，包含有步骤：

将信源图像序列划分成一种图像组集合，每个图像组中包含有一个此后称为 I 帧的第一帧，后随有至少一对此后称为 PB 帧的图像帧；

将每个 I 帧和 PB 帧划分成若干个在空间上不重叠的像素块；

对所述 I 帧中来的、此后称为 I 块的各像素块进行编码，诸 I 块与该图像组中任一其它帧之间没有相关性；

对于所述 PB 帧中时间上为第二帧内来的、此后称为 P 块的各像素块，根据先前之 I 帧中的 I 块或者先前之 PB 帧中的 P 块，推导出各个运动矢量和相应的预测值；

对于所述 PB 帧之第一帧中来的、此后称为 B 块的每个像素块，从具有与之相同位置的所述 P 块之运动矢量中推导出一个前向运动矢量，对每个 B 块可得到先前之 I 帧中一个相关联的 I 块或者先前之 PB 帧中一个相关联的 P 块，此后它们分别称为 If 块或 Pf 块；

对于所述 PB 帧之第一帧中来的每个 B 块，从具有与之相同位置的所述 P 块之运动矢量中推导出一个反向运动矢量，对每个 B 块可得到所述 PB 帧之 P 帧中一个相关联的 P 块，它此后称为 Pb 块；

对于每个 B 块的编码选择一种预测模式；

根据先前之 I 帧中的 I 块或者先前之 PB 帧中的 P 块对所述 PB 帧之第二帧中的各 P 块进行预测编码；

按照所选择定的预测模式对 B 块进行预测编码，

其中，编码每个 B 块时预测模式的选择中，对于每个 B 块包含有下面一系列步骤：

推导出该 B 块与一个像素块之间的绝对差值总和，该像素块内诸像素值是 Pb 块内诸像素值和 Pf 块或 If 块内诸像素值之平均值，此绝对差值总和此后称为 SADbidir（SAD 双向）；

推导出该 B 块与位置同该 B 块一样的 PB 帧之第二帧内的 P 块之间的绝对差值总和，它此后称为 SADb（SAD 反向）；

当 SADb 大于 SADbidir 时，根据具有与该 B 块相同位置的所述 P 块对该 B 块的预测编码模式作出选择；

当 SADb 小于 SADbidir 时：

推导出所述运动矢量与具有与该 B 块那样相同位置的所述 PB 帧之 P 帧中所述 P 块的预测值两者间的差值；

当所得到的差值小于一个预定阈值时，根据所述 PB 帧之第二帧内的 P 块、以及 I 块或先前之 PB 帧内的 P 块对该 B 块的预测编码作出选择；

当所得到的差值大于该预定阈值时，根据先前之 I 帧中的 I 块或者根据先前之 PB 帧中的 P 块推导出该 B 块之绝对差值总和的最小值，并根据该 I 块或先前之 PB 帧中的 P 块对该 B 块的预测编码模式作出选择。

2. 一种对一个图像序列进行编码的系统，它包含有：

一种装置，用于将信源图像序列划分成一种图像组集合，每个图像组中包含有一个此后称为 I 帧的第一帧，后随有至少一对此后称为 PB 帧的预测编码帧；

一种装置，用于将每个 I 帧和 PB 帧划分成若干个在空间上不重叠的像素块；

一个运动估值器，用于对所述 PB 帧中时间上为第二帧内的、此后称为 P 块的宏块，根据先前之 I 帧中的 I 块或者先前之 PB 帧中的 P 块推导出运动矢量以及相应的预测值，并用于对所述 PB 帧之第一帧中此后称为 B 块的每个宏块，从具有与之相同位置之 P 块的所述运动矢量中推导出一个前向运动矢量，这可以为每个 B 块得到在先前之 I 帧中一个相关联的 I 块，或者在先前之 PB 帧中一个相关联的 P 块，它们此后分别称为 If 块或 Pf 块，又用于从具有与之相同位置之 P 块的所述运动矢量中对每个 B 块推导出一个反向运动矢量，这可以为每个 B 块得到一个此后称为 Pb 块的相关联的 P 块；

用于在编码每个 B 块时选择出一个预测模式的装置；

用于对所述 I 帧中来的、此后称为 I 块的各宏块进行编码的装置，诸 I 块与该图像组中的任一其它帧之间没有相关性，并用于根据先前之 I 帧中的 I 块或者先前之 PB 帧中的 P 块对所述 PB 帧之第二帧中的 P 块进行预测编码，又用于对所述 PB 帧之第一帧中的 B 块进行预测编码，

其中，用于编码每个 B 块时选择出预测模式的装置，对每个 B 块

实施下面一系列步骤：

　　推导出该 B 块与一个像素块之间的绝对差值总和，该像素块内诸像素值是 Pb 块内诸像素值和 Pf 块或 If 块内诸像素之平均值，此绝对差值总和此后称为 SADbidir（SAD 双向）；

5　　推导出该 B 块与 Pb 块之间的绝对差值总和，它此后称为 SADb（SAD 反向）；

　　当 SADb 大于 SADbidir 时，根据具有与 B 块那样相同位置的所述 P 块对该 B 块的预测编码模式作出选择；

　　当 SADb 小于 SADbidir 时：

10　　　推导出所述运动矢量与具有与该 B 块那样相同位置的所述 PB 帧之 P 帧中所述 P 块的预测值两者间的差值；

　　　　当所得到的差值小于一个预定阈值时，根据该 P 块以及 I 块或先前之 PB 帧内的 P 块对该 B 块的预测编码模式作出选择；

　　　　当所得到的差值大于该预定阈值时，根据先前之 I 帧中的 I 块

15 或者根据先前之 PB 帧中的 P 块推导出该 B 块之绝对差值总和的最小值，并根据该 I 块或先前之 PB 帧中的 P 块对该 B 块的预测编码模式作出选择。

　　3．一种对一个图像序列进行编码的系统，它包含有：

　　用于将信源图像序列划分成一种图像组集合的装置，每个图像组中

20 包含有一个此后称为 I 帧的第一帧，后随有至少一对此后称为 PB 帧的预测编码帧；

　　用于将每个 I 帧和 PB 帧划分成若干个在空间上不重叠的像素块的装置；

　　一个运动估值器，用于对所述 PB 帧中时间上为第二帧内的、此后

25 称为 P 块的宏块，根据先前之 I 帧中的 I 块或者先前之 PB 帧中的 P 块推导出运动矢量以及相应的预测值，并用于对所述 PB 帧之第一帧中此后称为 B 块的每个宏块，从具有与之相同位置之 P 块的所述运动矢量中推导出一个前向运动矢量，这可以为每个 B 块得到在先前之 I 帧中一个相关联的 I 块，或者在先前之 PB 帧中一个相关联的 P 块，它们此后

30 分别称为 If 块或 Pf 块，又用于从具有与之相同位置之 P 块的所述运动矢量中对每个 B 块推导出一个反向运动矢量，这可以为每个 B 块得到

3

一个此后称为 Pb 块的相关联的 P 块；

用于在编码每个 B 块时选择出一个预测模式的装置；

用于对所述 I 帧中来的、此后称为 I 块的各宏块进行编码的装置，诸 I 块与该图像组中的任一其它帧之间没有相关性，并用于根据先前之 I 帧中的 I 块或者先前之 PB 帧中的 P 块来对所述 PB 帧之第二帧中的 P 块进行预测编码，又用于对所述 PB 帧之第一帧中的 B 块进行预测编码，

其中，所述编码系统中还包含一种可存储计算机程序的计算机可读介质，程序本身中包含一个指令集来取代至少某种所述装置，并且在一个计算机或数字处理器的控制下可以执行该程序，以便依靠对被取代装置所完成的那些相同功能的具体实施，来实现按照权利要求1中的编码方法。

说 明 书

选择 B 帧编码模式的视频编码方法

本发明涉及一种对信源图像序列进行编码的方法，包含有步骤：

5   将信源图像序列划分成一种图像组集合，每个图像组中包含有一个此后称为 I 帧的第一帧，后随有至少一对此后称为 PB 帧的图像帧；

将每个 I 帧和 PB 帧划分成若干个在空间上不重叠的像素块；

对所述 I 帧中来的、此后称为 I 块的各像素块进行编码，诸 I 块与该图像组中任一其它帧之间没有相关性；

10   对于所述 PB 帧中时间上为第二帧内来的、此后称为 P 块的各像素块，根据先前之 I 帧中的 I 块或者先前之 PB 帧中的 P 块，推导出各个运动矢量和相应的预测值；

对于所述 PB 帧之第一帧中来的、此后称为 B 块的每个像素块，从具有与其相同位置的所述 P 块之运动矢量中推导出一个前向运动矢

15   量，对每个 B 块可得到先前之 I 帧中一个相关联的 I 块或者先前之 PB 帧中一个相关联的 P 块，此后它们分别称为 If 块或 Pf 块；

对于所述 PB 帧之第一帧中来的每个 B 块，从具有与其相同位置的所述 P 块之运动矢量中推导出一个反向运动矢量，对每个 B 块可得到所述 PB 帧之 P 帧中一个相关联的 P 块，它此后称为 Pb 块；

20   对于每个 B 块的编码，选择一种预测模式；

根据先前之 I 帧中的 I 块或者先前之 PB 帧中的 P 块对所述 PB 帧之第二帧中的各 P 块进行预测编码；

按照所选择定的预测模式对 B 块进行预测编码。

本发明还涉及一种实现所述方法的系统。

25   例如，本发明可应用于甚低比特率的视频编码中。

ITU（国际电信联盟）制定的低比特率可视电话产品和技术方面的标准汇编于 ITU 标准 H.320 和 H.324 中。这些标准说明了能满足各种成分即音频、视频、复用器、控制协议和调制解调器等所要求的一切方面。H.320 专用于通过 ISDN（综合业务数据网）电话线的会议电视或可视电话中。

30   H.324 的应用目标是通过 GSTN（全球交换电话网）模拟电话线传输可视电话。这两个标准都支持 H.263 建议中的视频编码，

H.263 建议中说明了低比特率视频信号的压缩。H.263 建议中包含有视频编码器内的四种可选模式。这些可选模式之一称为 PB 帧模式，它给出了一种对 PB 帧进行编码的方法。H.263 建议的第二版本称为 H.263+，其开发在于改善图像质量，它包含一些新的可选方案。因此，一种对原来的 PB 帧模式作出改善的所谓改善型 PB 帧模式之方案，提供出了一种对 PB 帧进行编码的新方法。

一个图像帧序列可由一系列 I 帧和 PB 帧组成。图像中的 I 帧按照帧内模式编码，这意味着，I 帧利用图像内的空间冗余进行编码，它不参考另一帧图像。P 帧是根据先前之 P 图像或 I 图像进行预测编码的。因此，当编码一帧 P 图像时，除了象编码 I 图像那样利用空间冗余之外，还利用了该 P 图像与用作参考图像（基本上为先前的 I 图像或 P 图像）的先前图像之间的时间冗余。一帧 B 图像有两帧时间上先后的参考图像，它通常是根据一帧先前的 P 或 I 图像和当前在重建的 P 图像进行预测编码的。一个 PB 帧由相继的两帧图像构成，即一个首先的 B 帧和一个随后的 P 帧，它们编码成一个单元。

一种按照 PB 帧模式来对一个 PB 帧进行编码的方法示例于图 1 中。图中示明，一个 PB 帧由一个 B 帧的 B 和一个 P 帧的 P2 组成。该 B 帧的 B 之前后分别为一个先前的 P 图像 P1 和当前在重建的 P 图像 P2。本例子中示明的 P1 是一个 P 图像，但 P1 也可以是一个 I 图像，它用作为编码 P 图像 P2 和编码 B 图像 B 时的参考图像。在 PB 帧模式中，B 帧内的一个 B 块可以是按前向预测或按双向预测进行编码的。B 块按前向预测进行编码时，是基于先前的 I 或 P 图像 P1；B 块按双向预测进行编码时，是基于先前的 I 或 P 图像 P1 和当前在重建的 P 图像 P2 两者的。对于 PB 帧的 P 图像 P2 而言，是参考了图像 P1 来推导出一个运动矢量集合 MV 的。事实上，对于 P2 中的每个宏块，通过块匹配法使它与 P1 中的一个宏块相关联，并推导出一个相应的运动矢量 MV。B 块的运动矢量是从先前对 P1 得出的运动矢量集合中推导出的。所以，对于一个 B 块可如下地计算其前向运动矢量 MVf 和反向运动矢量 MVb：

$$MVf = (TRb \times MV) / TRd$$
$$MVb = ((TRb - TRd) \times MV) / TRd \quad (1)$$
$$MVb = MVf - MV$$

2

式中，TRb 是在先前的 P 帧 P1 为准下 B 图像在时间基准上的增量值，TRd 是在先前的 I 或 P 图像 P1 为准下当前的 P 帧 P2 在时间基准上的增量值。

考虑图 1 上 B 图像内的一个宏块 AB。此宏块 AB 的位置与先前重建的 P2 内宏块 A2B2 即 Prec 的位置相同。前向运动矢量 MV 与属于 P1 之宏块 A1B1 来的宏块 A2B2 相关联。与宏块 AB 相关联的前向运动矢量 MVf 和反向运动矢量 MVb 两者是按式（1）中所示自 MV 中导出的。如图 1 上所示，通过前向运动矢量 MVf 和反向运动矢量 MVb 而与宏块 AB 相关联的 P1 和 P2 内的宏块分别是 K1M1 和 K2M2。

对于双向预测与前向预测之间的选择，是在 B 图像中的宏块层级上作出的，它取决于 MVb 点的位置所在。具体地，在 B 块上 AB 里的 MB 部分处，因 MVb 点位于 Prec 之内，该 MB 部分为双向预测，B 块上该 MB 部分的预测公式为：

$$MB\,(i,j) = [\,A1M1\,(i,j) + A2M2\,(i,j)\,]\,/\,2$$

式中，i 和 j 为各像素的空间座标值。

另外，在 B 块上 AB 里的 AM 部分处，因 MVb 点位于 Prec 之外，该 AM 部分为前向预测，B 块上该 AM 部分的预测公式为：

$$AM\,(i,j) = K1A1\,(i,j)$$

按照 PB 帧模式中的一种改善型编码 PB 帧的方法，说明于欧洲专利申请 EP 0782343A2 中。它公开了一种在双向预测帧中编码各宏块时的预测方法。该方法介绍了一种增量运动矢量，针对导出的前向和反向运动矢量分别用增量运动矢量作相加或相减。当图像序列中的运动不为线性形态时，所介绍的方法或许是适当的；然而，对于发生镜头切换的图像序列，该方法是完全不合适的。确实，当先前的 P 帧与一个 PB 帧中的 B 帧部分之间有镜头切换时，双向预测和前向预测将给出错误的编码。除了增量运动矢量在实现上会十分加重 CPU 的负担外，还将导致不必要的耗费和复杂的计算。

本发明的一个目的是改进现有编码方法的效率，同时降低 CPU 的负担。更准确地，本发明的一个目的是提供一种高效的策略或方法，它对于一个给定的 B 帧宏块之编码可给出最合适的预测模式选择。

因此，编码每个 B 块时在预测模式的选择中，对每个 B 块包含有

下面一系列步骤：

推导出该 B 块与一个像素块之间的绝对差值总和，该像素块内诸像素值是 Pb 块内诸像素值和 Pf 块或 If 块内诸像素之平均值，此绝对差值总和此后称为 SADbidir（SAD 双向）；

推导出该 B 块与位置同该 B 块一样的 PB 帧之第二帧内的 P 块之间的绝对差值总和，它此后称为 SADb（SAD 反向）；

当 SADb 大于 SADbidir 时，根据具有与 B 块相同位置的所述 P 块来对 B 块的预测编码模式作出选择；

当 SADb 小于 SADbidir 时：

推导出所述运动矢量与具有与该 B 块那样相同位置的所述 PB 帧之 P 帧中所述 P 块的预测值两者间的差值；

当所得到的差值小于一个预定阈值时，根据所述 PB 帧之第二帧内的 P 块、以及 I 块或先前之 PB 帧内的 P 块来对 B 块的预测编码模式作出选择；

当所得到的差值大于该预定阈值时，根据先前之 I 帧中的 I 块或者根据先前之 PB 帧中的 P 块推导出该 B 块之绝对差值总和的最小值，并根据该 I 块或先前之 PB 帧中的 P 块来对 B 块的预测编码模式作出选择。

对于 B 块的编码，本权利要求之方法给出了一种预测模式的选择策略，以供在前向模式、反向模式和双向模式中作选择应用。该选择是基于 SAD（绝对差值总和）计算和运动矢量相关性的。选择策略依据于一种特定的规则，它比较三种预测模式的 SAD 值，并引入运动矢量相关性。此外，对于编码 B 帧时的预测模式选择，是在任何 P 或 B 图像编码之前作出的。因此，由于所提出的选择策略是在原来的图像上实施的，所以在 SAD 计算、特别是 SADbidir 计算中，不需要进行耗费 CPU 能力的、对 B 帧的预先双向预测。所提出之方法的主要优点不仅有利于双向预测，并在无运动状态时也能实现反向预测。因此，本方法对于一个给定的 B 帧宏块，可导得一种合适的预测模式选择。

本发明的一个优选实施例中，按照本发明之一种方法，或是可以由一种电子电路所构成的硬体系统来实现，它能完成所建议之方法中的各个步骤，或是可以部分地依靠计算机可读出介质中存储的一个指令集来

4

完成。

现在，参考后面说明的实施例并结合考虑各个附图，对本发明的具体情况作出阐述，各个附图中，

图1示例出按照PB帧模式的一种先有技术解码方法；

图2示明一个用于编码的图像序列；

图3是编码系统中由各部分组成的一个方框图；

图4表明对运动矢量预测值是如何确定的；

图5是一个流程图，示明按照本发明对B块进行编码的各个步骤，它们导向预测模式的选择。

下面各节段中可能发生对名词"块"的误解。当读到"块"字时，应理解为它是ITU标准中所定义的"宏块"。

图2示明一个信源图像帧序列，它们必须按照本发明之方法进行编码。这个所示的序列由第一个I帧I0以及在时间上后随的一系列PB帧组成。每个PB帧PB1、PB2、PB3由称为B帧的一个第一帧和称为P帧的一个第二帧构成。因此，PB1由B帧B1和随后的P帧P2组成，PB2由B帧B3和随后的P帧P4组成，PB3由B帧B5和随后的P帧P6组成，……。

各个帧将以下面给出的次序编码。I0按照帧内模式首先编码，也即编码中不参考任何其它图像帧。然后，P2参考I0进行预测编码；接着，B1参考I0和P1进行编码，B1是在编码器之内于内部重建出来的；再后，P4参考P2进行编码，接着，P3参考P2和P4进行编码，B3也是于内部重建出来的。因此，在一个PB帧的B块之前使序列中该PB帧的每个P块先传输，并参考先前的I或P图像进行编码。在相应的该PB帧的P图像之后再对每个B图像进行编码，这是参考所述相应之PB帧中的P图像和先前已编码的I或P图像进行编码的。

图2中提出的图像序列决不意味着是对可以按照本发明之一种方法进行编码的图像序列种类的一个限制。事实上，图像序列中也可以在两个P帧之间包含两个或多个接连的B帧。此种场合下，各个B帧按它们传输中的同样次序参考先前的I或P帧（它们先前已经编码好）以及次后一个P帧（它是当前重建出的）进行编码。

诸如图2中示明的一种图像序列是经由图3之系统内各级编码措施

5

逐帧图像地通过的，所述系统可以实现一种按照本发明之方法。首先，为了编码方便，由一个电路 DIV（I，P，B）将每个传输的帧划分成在空间上不重叠的 N×M 个像素（比如说 16×16 个像素）组成的诸多宏块。I、P 和 B 帧是不以同样的方式编码的，所以，它们通过系统时不遵循相同的信号路径。每一种类的各帧沿着一条它们所适应的路径前进。

编码时不需要参考任何其它图像的 I 帧，是从电路 DIV（I，P，B）直通至电路 DCT/Q 上的。这个电路 DCT/Q 将空间域内接收到的一个帧变换成频率域内的一个帧。它对划分成像素块的图像实施离散余弦变换，得到一个变换系数集，然后对该系数集进行量化。接着，自 DCT/Q 电路给出的已量化的系数通过到电路 COD 上进一步编码，并同时通过到电路 $IDCT/Q^{-1}$ 上。电路 $IDCT/Q^{-1}$ 先实施反量化，再通过反离散余弦变换将反量化的系数变换回到空间域中。由电路 REC（P）重建出 I 帧的每个像素块，然后将 I 图像存储入电路 MV（P）中的一个存储器部分里。

对于 P 帧，在由 DIV（I，P，B）将它划分成像素块后，传输至一个运动估值器 MV（P）上。MV（P）中包含一个存储器部分，在那里又存储入了先前已传输来的 I 图像或 P 图像。对于 P 图像内此后称为 P 块的每个宏块，参考当前已存储的先前之 I 或 P 图像来推导出一个运动矢量 MV。此运动矢量 MV 可以对原来诸宏块而不是重建出的诸宏块计算出其最小化的函数值 SAD（绝对差值总和）来推导得到。SAD 值给出于下面：

$$\text{SAD} = \sum_{m=1}^{16} \sum_{n=1}^{16} \left| B_{i,j}(m,n) - B_{i-u,j-v}(m,n) \right|$$

式中，$B_{i,j}(m,n)$ 表示在空间位置 (i, j) 上 16×16 像素之 P 块中的第 (m, n) 个像素，$B_{i-u,j-v}(m,n)$ 表示在空间位置 (i, j) 受矢量 (u, v) 偏移后的位置上先前之 I 或 P 图像中一个备选宏块内的第 (m, n) 个像素。运动矢量便是给出最小 SAD 值的 P 块与该备选宏块之间的偏移量。同时，在电路 MV（P）中，对每个运动矢量 MV 推导出一个关联的预测值 MVpred。推导出 MVpred 的一种可能方法由 H. 263 建议给出，如

图 4 中所示例，它示明了一个 P 块以及其相邻的邻接块。MVpred 定义为 MV1、MV2、MV3 的中间值，其中，MV1 是该 MV 左前方之宏块的运动矢量，MV2 是该 MV 正上方之宏块的运动矢量，MV3 是该 MV 右上方之宏块的运动矢量。除了对一个给定的 P 块推导出 MV 和 MVpred 之外，还按照在式（1）中给出的公式推导出一个前向运动矢量 MVf 和一个反向运动矢量 MVb。这两个运动矢量与位置同该 P 块一样的 PB 帧之 B 帧中的一个 B 块相关联。此外，与这个 B 块相关联 MVf 确定了一个先前之 I 或 P 帧内、此后称为 If 块或 Pf 块的 I 或 P 块。类似地，与这个 B 块相关联的 MVb 确定了一个该 PB 帧之 P 帧内、此后称为 Pb 块的 P 块。在这个已运动补偿的 P 帧与 MV（P）之存储器部分中存储的先前之 I 或 P 帧之间，通过有分支的加法器 S 来实施相减，将差值传输至 DCT/Q 运算单元上，形成一个经 DCT 变换和量化了的系数帧。然后，这个系数帧通过到 COD 单元上进一步编码，并同时通过到 IDCT/Q^{-1} 和 REC（P）单元上。这里，REC（P）根据从电路 IDCT/Q^{-1} 中接收到的差值帧、从运动估值器 MV（P）中接收到的运动矢量、以及在 MV（P）之存储器部分中已存储的先前之 I 或 P 帧这三者的关联性，重建出该 P 帧中的每一个宏块。在重建之后，MV（P）中的存储器部分以当前的 P 帧数据进行更新。

B 帧直接通过到一个预测器 PRED（B）上，用以按照前向、反向或双向预测模式进行预测编码。当于 PRED（B）内进行编码时，在加法器 S 上从初始宏块中减去已运动补偿的宏块，差值通过 DCT/Q，然后到达 COD 进一步编码。对于 B 帧的每个宏块，必须在三种可能的预测模式中作出一个选择。事实上，在有效地开始编码 B 图像之前，先对所有宏块顺序地作出这些预测模式的选择。这些选择是根据从 MV（P）中接收到的关于原先的 I 或 P 图像以及 PB 帧中原来的 P 图像之数据作出的，两类图像数据都存储于 MV（P）之存储器部分中。PB 帧中的 P 图像是当前在 REC（P）中重建的，并且与 PB 帧中的 B 图像作为一个单元进行编码。确实，在 REC（P）中重建一个 P 块之后，B 帧中相同位置上的 B 块也按照对该宏块先前已选择的预测模式在 PRED（B）中进行预测编码。

对于 B 帧的一个宏块来说，按照本发明之策略来作出预测模式选择
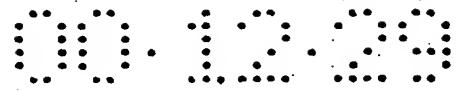
的过程示明于图 5 的流程图中。人们可回想到，与每个 B 块相关联的几个运动矢量，是先前在 MV（P）中计算出的下列运动矢量：PB 帧之 P 帧中具有象 B 块那样相同位置的 P 块之运动矢量 MV 及其预测值 MVpred；确定出 Pf 块的前向运动矢量 MVf；以及确定出 Ib 或 Pb 块的反向运动矢量 MVb。

对于传输至 PRED（B）上的一个 B 帧之每个宏块 Mbck[n] 来说，步骤 1 中，在 Mbck[n] 与一个宏块之间推导出 SAD 值，该宏块具有的像素值是 PB 帧之 P 帧中相关 Pb 块的像素值以及先前的 I 或 P 帧中 If 或 Pf 块的像素值之平均值。此后，这个 SAD 称为 SADbidir。同时，步骤 2 中，在 Mbck[n] 与该 Pb 块之间推导出 SAD 值，并称之为 SADb。随后，步骤 3 中，将 SADbidir 与 SADb 进行比较，在 SADbidir 大于 SADb 的场合下，对 Mbck[n] 作出反向预测模式的选择。这意味着，在对所有 B 块作出选择之后，Mbck[n] 将参考 PB 帧中的 P 图像进行预测编码。如图 5 中的反馈连接线所示，现在提供出一个新的 B 块，需对它作出预测模式选择。

另一种场合下，也即当 SADbidir 小于 SADb 时，便实施一项运动矢量相关性测试。在步骤 4 和 5 中，将运动矢量 MV 与其预测值 MVpred 进行比较，它们是分别与 PB 帧之 P 帧中具有 B 块那样相同位置的 P 块相关联的，以及与先前 MV（P）中计算出的值相关联的。当 MV－MVpred 之差值小于一个预定阈值 t1 时，在步骤 7 中选择双向预测模式。编码 B 帧时，随后需参考先前的 I 或 P 图像以及 PB 帧中的 P 图像对 B 块 Mbck[n] 进行预测编码。当 M－MVpred 之差值大于 t1 时，在步骤 8 中选择前向预测模式。此种场合下，需参考先前的 I 或 P 图像对 B 块 Mbck[n] 进行预测编码。然后，提供出一个新的 B 帧宏块。按照前向预测模式编码一个 B 块时，在步骤 6 中实施一种前向运动估值 MVF。它的要点是参考先前的 I 或 P 图像对 B 块推导出一个前向运动矢量，推导中借助于对参考先前的 I 或 P 图像计算出的 B 块使其绝对差值总和 SAD 最小化。

应当指出，关于所说明的编码方法和系统，可对之提出修正或改进而偏离不开本发明的范畴。例如，很显然，此种编码方法可以用几种方式来实现，诸如依靠电子电路硬件，或者，另一种方式是依靠在计算机

可读介质中存储入一个指令集，所述指令集可取代至少一部分所述电路，并能在一个计算机或数字处理器的控制下予以执行，以便实现如所述被取代电路中可完成的同样功能。此种场合下，这些指令在一个计算机程序中组合于一起，该程序可装载和存储于所述介质中，使得任何编码系统，诸如上面叙述的并包括所述介质的编码系统，依靠实施被取代电路可完成的那些同样的功能来实现所说明的编码方法。
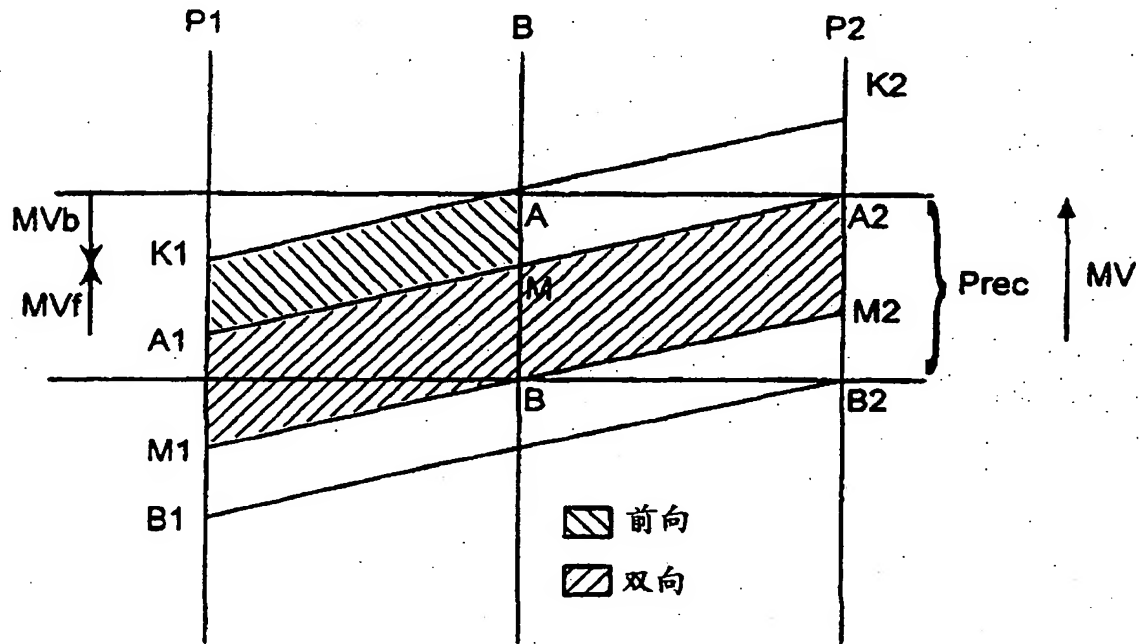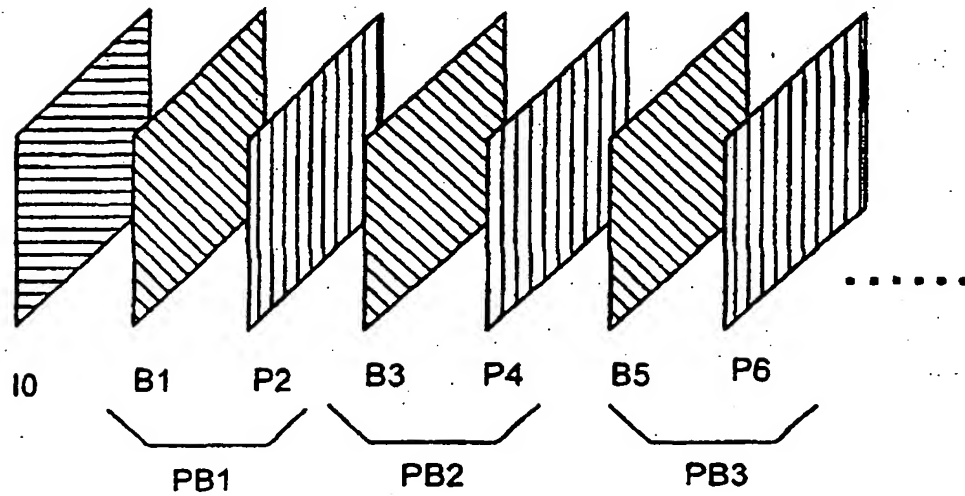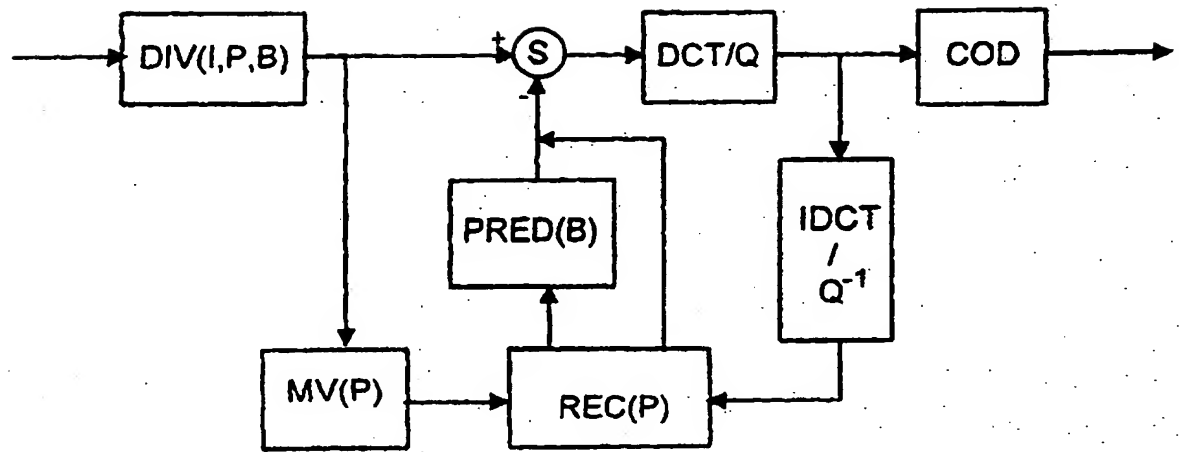
说 明 书 附 图

图 1

图 2

图 3

图 4

2

```
                          Mbck[n]
                    ┌────────┴────────┐
                    │                 │
                ┌───▼───┐ 1       2 ┌─▼─────┐
                │SADbidir│         │ SADb  │
                └───┬───┘          └───┬───┘
                    └────────┬─────────┘
                             │
                          ◇   ◇  N
                       Y ◇  <  ◇────────────┐
                          ◇   ◇             │
                           ╲ ╱  3           │
                            │               │
                      ┌─────▼─────┐ 4       │
                      │ MV-MVpred │         │
                      └─────┬─────┘         │
                            │               │
                          ◇  ◇  5           │
                         ◇ <t1 ◇  N         │
                          ◇  ◇──────┐       │
                           ╲╱       │       │
                         Y │        │       │
                           │    ┌───▼───┐ 6 │
                           │    │  MVF  │   │
                           │    └───┬───┘   │
                           │        │       │
                    7      │    8   │   9   │
                ┌──────┐   │ ┌──────▼┐  ┌───▼──┐
                │ 双向 │◄──┘ │ 前向  │  │ 反向 │
                └──┬───┘     └───┬───┘  └──┬───┘
                   │             │         │
                   ▼             ▼         ▼
```
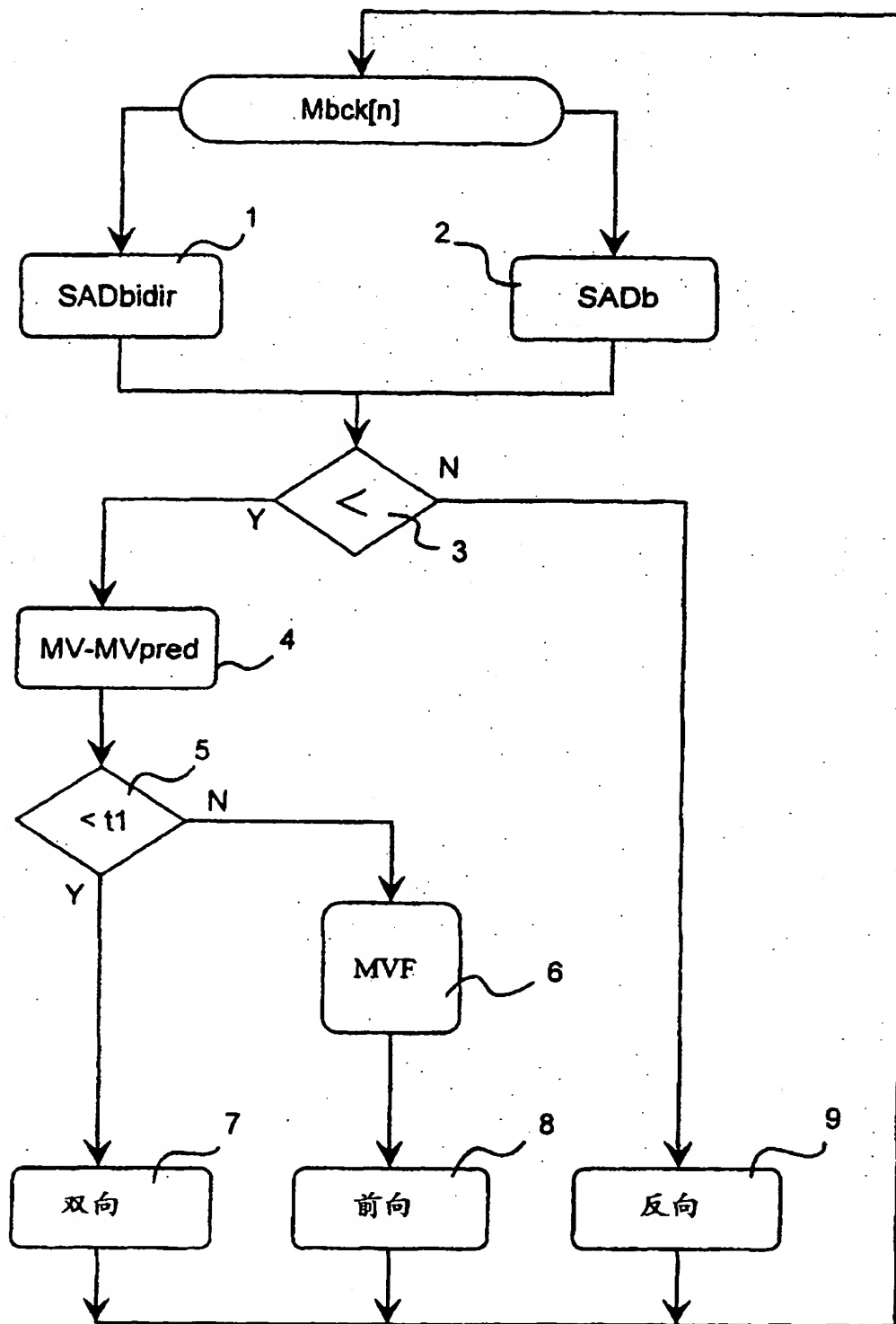
图 5

Video Encoding Method And System

FIELD OF THE INVENTION

The invention relates to a method of encoding a source sequence of pictures comprising the steps of: dividing a source sequence into a set of group of pictures, each group of pictures comprising a first frame, hereafter referred to as I-frame, followed by at least a pair of frames, hereafter referred to as PB-frames; dividing each I-frame and PB-frame into spatially non-overlapping blocks of pixels; encoding the blocks from said I-frame, hereafter referred to as I-blocks, independently from any other frame in the group of pictures; deriving motion vectors and corresponding predictors for the blocks from the temporally second frame of said PB-frame, hereafter referred to as the P-blocks, based on the I-blocks in the previous I-frame or the P-blocks in the previous PB-frame; deriving for each block from the first frame of said PB-frame, hereafter referred to as a B-block, a forward motion vector from said motion vector of the P-block with same location, allowing to obtain for each B-block an associated I-block in the previous I-frame or an associated P-block in the previous PB-frame, hereafter referred to as If-block or Pf-block, respectively; deriving for each B-block of the first frame of said PB-frame, a backward motion vector from said motion vector of the P-block with same location, allowing to obtain for each B-block an associated P-block in the P-frame of said PB-frame, hereafter referred to as the Pb-block; choosing a prediction mode for the encoding of each B-block; predictively encoding the P-blocks of the second frame of said PB-frame based on the I-blocks in the previous I-frame or the P-blocks in the previous PB-frame; predictively encoding the B-blocks following the chosen prediction mode.

The invention also relates to a system for carrying out said method.

The invention may be used, for example, in video coding at very low bit rate.

BACKGROUND ART

Standardization of low bitrate video telephony products and technology by the ITU (International Telecommunication Union) are compiled in the standards H.320 and H.324. These standards describe all the requirements to be satisfied for the different components audio, video, multiplexer, control protocol and modem. H.320 is dedicated to videoconferencing or videophony over ISDN (Integrated Services Data Network) phone lines. H.324 is aimed at videophony over GSTN (Global Switch Telephonic Network) analog phone lines. The two standards both support for video-coding the Recommendation H.263, which describes compression of low bitrate video signals. The H.263 Recommendation comprises four optional modes for a video coder. One of these optional modes is called the PB-frames mode, which gives a way of encoding a PB-frame. A second version of the H.263 Recommendation, called H.263+, was developed to improve the image quality and comprises some new options. Thus, an option called Improved PB-frames mode, which is an improvement of the original PB-frames mode, provides a new way of encoding a PB-frame.

A sequence of picture frames may be composed of a series of I-frames and PB-frames. A I-frame consists in a picture coded according to an Intra mode, which means that a I-frame is coded using spatial redundancy within the picture without any reference to another picture. A P-frame is predictively encoded from a previous P or I-picture. Thus, when coding a P-picture, temporal redundancy between the P-picture and a previous picture used as a picture reference, which is mostly the previous I or P-picture, is used in addition to the spatial redundancy as for a I-picture. A B-picture has two temporal references and is usually predictively encoded from a previous P or I-picture and the P-picture currently being reconstructed. A PB-frame consists of two successive pictures, a first B-frame and a subsequent P-frame, coded as one unit.

A method of coding a PB-frame in accordance to the PB-frames mode is illustrated in FIG. 1. It is shown a PB-frame composed of a B-frame B and a P-frame P2. The B-frame B is surrounded by a previous P-picture P1 and the P-picture P2 currently being reconstructed. It is shown in this example a P-picture P1, P1 may also be a I-picture and serves as a picture reference for the encoding of the P-picture P2 and the B-picture B. A B-block from the B-frame, in the PB-frames mode, can be forward or bidirectionally predictively encoded. The encoding of a B-block being forward predicted is based on the previous I or P-picture P1 and the encoding of a B-block being bidirectionally predicted is based on both the previous I or P-picture P1 and the P-picture P2 currently being reconstructed. A set of motion vectors MV is derived for the P-picture P2 of the PB-frame with reference to the picture P1. In fact for each macroblock of P2, a macroblock of P1 is associated by block matching and a corresponding motion vector MV is derived. Motion vectors for the B-block are derived from the set of motion vectors derived previously for P1. Therefore a forward motion vector MVf and a backward motion vector MVb are calculated for a B-block as follows:

where TRb is the increment in the temporal reference of the B-picture from the previous P-frame P1, and TRd is the increment in the temporal reference of the current P-frame P2 from the previous I or P-picture P1.

It is considered on FIG. 1 a macroblock AB of the B-picture. This macroblock AB has the same location as a macroblock A2B2, Prec, of P2 that was previously reconstructed. A forward motion vector MV is associated to the macroblock A2B2 from a macroblock A1B1, which belongs to P1. A forward motion vector MVf and a backward motion vector MVb, both associated with AB, are derived from MV as shown in (1). The macroblocks of P1 and P2 associated with the AB macroblock by the forward vector MVf and by the backward vector MVb are respectively K1M1 and K2M2 as illustrated on FIG. 1.

The choice between bidirectional prediction and forward prediction is made at the block level in the B-picture and it depends on where MVb points. Then a MB part of the B-block AB, for which MVb points inside Prec, is bidirectionally predicted and the prediction for this part of the B-block is: $MB(ij)=[A1M1(ij)+A2M2(ij)]/2$ where i and j are the spatial coordinates of the pixels.

A AM part of the B-block AB, for which MVb points outside Prec, is forward predicted and the prediction for this part of the B-block AB is: $AM(i,j)=K1A1(i,j)$.

2

An improved method of encoding a PB-frame according to the PB-frames mode is described in the European Patent Application EP 0 782 343 A2. It discloses a predictive method of coding the blocks in the bidirectionally predicted frame, which method introduces a delta motion vector added to or subtracted from the derived forward and backward motion vectors respectively. The described method may be relevant when the motion in a sequence of pictures is non-linear, however it is totally unsuitable for a sequence of pictures where scene-cuts occur. Indeed, when there is a scene cut between a previous P-frame and the B-part of a PB-frame, bidirectional and forward prediction give an erroneous coding. Besides the implementation of the delta vector, which is costly in terms of CPU burden, may result in unnecessary expensive and complicated calculations.

SUMMARY OF THE INVENTION

It is an object of the invention to ameliorate efficiency of existing coding methods, while decreasing CPU burden, and more precisely an object of the invention is to provide an efficient strategy or method which permits to make the most suitable choice among prediction modes for the coding of a given macroblock of a B-frame.

Thus, the choice of the prediction mode for the encoding of each B-block comprises for each B-block in series the steps of: deriving the sum of absolute difference between the B-block and a block with pixels values being the means of the pixels values of the Pb-block and of the Pf-block or If-block, hereafter referred to as SADbidir; deriving the sum of absolute difference between the B-block and the P-block in the second frame of the PB-frame with same location as the B-block, hereafter referred to as SADb; when SADb is greater than SADbidir, making the choice of predictively encoding the B-block based on said P-block with same location as the B-block; when SADb is lower than SADbidir: deriving the difference between said motion vector and said predictor of the P-block in the P-frame of said PB-frame with same location as the B-block; when the obtained difference is lower than a predetermined threshold, making the choice of predictively encoding the B-block based on the P-blocks of the second frame of said PB-frame and the I-blocks or the P-blocks in the previous PB-frame; when the obtained difference is greater than the predetermined threshold, deriving the minimum of the sum of absolute difference for the B-block based on the I-blocks in the previous I-frame or on the P-blocks in the previous PB-frame, and making the choice of predictively encoding the B-block based on the I-blocks or the P-blocks in the previous PB-frame.

The claimed method gives, for the coding of a B-block, a strategy for the choice of the prediction mode to be used among the forward, backward and bidirectional modes. The choice is based upon SAD (Sum of Absolute Difference) calculation and motion vector coherence. The strategy is based upon a specific order in the comparisons of the SAD values for the three prediction modes and the introduction of motion coherence. Furthermore the choice of the prediction mode for the encoding of a B-frame is made before any P or B-picture is encoded. Thus, because the proposed strategy is performed on original pictures, the SAD calculations, particularly the calculation of SADbidirectional, do not require a prior bidirectional prediction for the B-frame, which is CPU

consuming. The proposed method has the main advantage of not being in favor of bidirectional prediction and allows to perform backward prediction when there is no motion. Thus, the method leads to a suitable choice of prediction mode for a given block of a B-frame.

In a preferred embodiment of the invention, a method according to the invention may either be carried out by a system constituted of wired electronic circuits that may perform the various steps of the proposed method. This method may also be partly performed by means of a set of instructions stored in a computer-readable medium.

BRIEF DESCRIPTION OF THE DRAWINGS

The particular aspects of the invention will now be explained with reference to the embodiments described hereinafter and considered in connection with the accompanying drawings, in which:

FIG. 1 illustrates a prior art decoding method according to the PB-frames mode;

FIG. 2 shows a sequence of pictures for encoding;

FIG. 3 is a block diagram of the various steps of a coding system;

FIG. 4 allows to understand how the predictor of a motion vector is defined;

FIG. 5 is a block diagram of the various steps in the encoding of a B-block leading to the choice of a prediction mode in accordance with the invention.

DETAILED DESCRIPTION OF THE INVENTION

A misuse of the word "block" may occur in the following paragraphs. When reading block one must understand macroblock, as defined in ITU standards.

FIG. 2 depicts a source sequence of picture frames that has to be encoded following a method in accordance with the invention. This shown sequence is organized in a first I-frame Io temporally followed by a series of PB-frames. Each PB-frame PB1, PB2, PB3 is constituted of a first frame, say, a B-frame and a second frame, say a P-frame. Thus, PB1 is composed of a B-frame B1 and a subsequent P-frame P2, PB2 is composed of a B-frame B3 and a subsequent P-frame P4, PB3 is composed of a B-frame B5 and a subsequent P-frame P6 . . . .

The various frames will be encoded in the order given hereinafter. Io is first encoded according to an Intra mode, i.e. without reference to any other picture. P2 is, then, predictively encoded in reference to Io and, subsequently, B1 is encoded in reference to Io and P1, which is, inside the encoder, internally reconstructed. P4 is then encoded in reference to P2 and, subsequently, B3 is encoded in reference to P2 and P4, which is internally reconstructed too. Thus, each P-block of a PB-frame in the sequence is transmitted and encoded before a B-block of the PB-frame and in reference to the previous I or P-picture. Each B-picture is encoded after the corresponding

4

P-picture of the PB-frame and in reference to said corresponding P-picture of the PB-frame and to the previous I or P-picture encoded.

The sequence of pictures proposed in FIG. 2 is by no means a limitation of the sort of sequences of pictures, which may be encoded following a method in accordance with the invention. In fact, the sequence may also comprise two or more successive B-frames between two P-frames. In such case the B-frames are encoded in the same order as they are transmitted in reference to the previous I or P-frame and the next P-frame, which was previously encoded and which is currently reconstructed.

A sequence of pictures, such as the one described in FIG. 2, is passed picture-by-picture through the various coding steps of the system in FIG. 3, said system being provided for carrying a method in accordance to the invention. First a circuit DIV(I,P,B) divides each transmitted frame into spatially non-overlapping N.times.M, say 16.times.16, macroblocks of pixels for encoding convenience. I, P and B frames are not encoded in the same way, so, they do not follow the same path through the system. Each sort of frame follows an adapted path.

A I-frame, whose encoding does not require reference to any other picture, is passed from the circuit DIV(I,P,B) to a circuit DCT/Q. This circuit DCT/Q transforms a frame received in the spatial domain into a frame in the frequency domain. It applies a discrete cosine transform to the picture divided into blocks of pixels, resulting in a set of transform coefficients, which are then quantized. These quantized coefficients, issuing from the DCT/Q circuit are, then, passed to a circuit COD for further encoding and at the same time to a circuit IDCT/Q-1. The circuit IDCT/Q-1 dequantizes and transforms the coefficients by inverse discrete cosine transform, back in the spatial domain. A circuit REC(P) reconstructs each block of the I-frame and then the I-picture is stored in a memory part of a circuit MV(P).

A P-frame, after division into blocks of pixels by DIV(I,P,B), is transmitted to a motion estimator MV(P). MV(P) comprises a memory part where the previous transmitted I or P-picture is stored. A motion vector MV is derived for each block of the P-picture, hereafter referred to as P-block, in reference to the previous I or P-picture currently stored. This vector MV may be possibly derived by minimizing a function SAD (Sum of Absolute Difference) computed on original blocks rather than reconstructed ones, which is given hereinbelow: ##EQU1##

where $B_{i,j}(m,n)$ represents the (m,n)th pixel of the 16.times.16 P-block at the spatial location (ij) and $B_{i-u,j-v}(m,n)$ represents the (m,n)th pixel of a candidate macroblock in the previous I or P-picture at the spatial location (ij) displaced by the vector (u,v). The motion vector is the displacement between the P-block and the candidate macroblock giving the smallest SAD. Simultaneously, in the circuit MV(P), an associated predictor Mvpred is derived for each motion vector Mv. A possible way of deriving MVpred is given by the H.263 Recommendation as illustrated in FIG. 4, which depicts a P-block and its adjacent neighbouring blocks. MVpred is defined as the median value of MV1, MV2, MV3 where MV1 is the motion vector of the previous macroblock, MV2 is the motion vector of the above macroblock and MV3 is the motion vector of the above right macroblock. In addition to the derivation of MV and MVpred for a given P-block,

a forward motion vector MVf and a backward motion vector MVb are derived following the equations given in set 1. These two motion vectors are associated to a B-block in the B-frame of the PB-frame with same location as the P-block. Furthermore MVf, associated to this B-block, defines a I or P-block in the previous I or P-frame, hereafter referred to as If or Pf-block. Similarly MVb, associated to this B-block, defines a P-block in the P-frame of the PB-frame, hereafter referred to as a Pb-block. The difference between this motion compensated P-frame and the previous I or P-frame stored in the memory part of MV(P) is performed in the tap adder S and transmitted to the unit DCT/Q resulting in a quantized transformed frame. This one is, then, passed to the unit COD for further encoding and, at the same time, to the units IDCT/Q-1 and REC(P). Here, REC(P) reconstructs each block of the P-frame from the association of the differential frame received from the circuit IDCT/Q-1, the motion vectors received from the motion estimator MV(P) and the previously I or P-frame stored in the memory part of MV(P). After being reconstructed the memory part of Mv(P) is updated with the current P-frame.

A B-frame is passed directly to a predictor PRED(B) for being predictively encoded according to a forward, backward or bidirectional prediction mode. When encoded in PRED(B)the motion compensated block is subtracted in S from the initial block, the difference is passed through DCT/Q and then to COD for further encoding. A choice has to be made among the three possible prediction modes for each macroblock of the B-frame. In fact, before the effective encoding of the B-picture starts, these choices of the prediction modes are made in series for all the macroblocks. These choices are based upon data received from MV(P) about the original previous I or P-picture and the original P-picture of the PB-frame, both pictures being stored in the memory part of MV(P). The P-picture of the PB-frame is currently reconstructed in REC(P) and it is coded with the B-picture of the PB-frame as one unit. Indeed, once a P-block is reconstructed in REC(P), the B-block of same location in the B-frame is also, in PRED(B), predictively encoded following the prediction mode that was previously chosen for this macroblock.

A strategy in accordance with the invention leading to the choice of the prediction mode for a macroblock of a B-frame is depicted in the diagram of FIG. 5. One must recall that in association to each B-block several motion vectors were previously calculated in MV(P): the motion vector MV and its predictor MVpred of the P-block in the P-frame of the PB-frame with same location as the B-block, the forward motion vector MVf that defines the Pf-block and the backward motion vector MVb that defines the Ib or Pb-block.

For each macroblock Mbck[n] of a B-frame transmitted to PRED(B), the SAD is, in a step 1, derived between Mbck[n] and a block with pixels values being the means of the pixels values of the associated Pb-block of the P-frame of the PB-frame and of the If or Pf-block of the previous I or P-frame. This SAD is hereafter referred to as SADbidir. Simultaneously, in a step 2, the SAD between Mbck[n] and the Pb-block is derived and referred to as SADb. Afterwards, SADbidir and SADb are compared in a step 3 and in the case of SADbidir being greater than SADb the choice is made of the backward prediction mode for Mbck[n], which signifies that once the choices are made for all the B-blocks, Mbck[n] will be predictively encoded in reference to the P-picture of the PB-frame. A new B-block, for which the choice of the prediction mode has to be made, is now provided, as shown with the feedback connection in FIG. 5.

In the other case, i.e. when SADbidir is lower than SADb, a motion vector coherence test is performed. The motion vector MV and its predictor MVpred, that are associated to the P-block with same location as the B-block in the P-frame of the PB-frame and that were previously calculated in MV(P), are compared in steps 4 and 5. When the difference MV-MVpred is lower than a predefined threshold t1, bidirectional prediction is chosen in a step 7. The B-block Mbck[n] will be later on, in the encoding of the B-frame, predictively encoded in reference to the previous I or P-picture and the P-picture of the PB-frame. When the difference is greater than t1, the forward prediction is chosen in a step 8 and the B-block Mbck[n] will be, in this case, predictively encoded in reference to the previous I or P-picture. A new macroblock of the B-frame is then provided. For the encoding of a B-block following a forward prediction mode, a forward motion estimation MVF is performed in a step 6. It consists in the derivation of a forward motion vector for the B-block in reference to the previous I or P-picture by minimizing the sum of absolute difference for the B-block calculated in reference to the previous I or P-picture.

It is to be noted that, with respect to the described coding method and system, modifications or improvements may be proposed without departing from the scope of the invention. For instance, it is clear that this coding method can be implemented in several manners, such as by means of wired electronic circuits or, alternatively, by means of a set of instructions stored in a computer-readable medium, said instructions replacing at least a part of said circuits and being executable under the control of a computer or a digital processor in order to carry out the same functions as fulfilled in said replaced circuits. In such a case, these instructions are assembled together in a computer program that can be loaded and stored in said medium and causes any encoding system, such as described above and including said medium, to be able to carry out, by means of an implementation of the same functions as those fulfilled by the replaced circuits, the described encoding method.

What is claimed is:

1. A method of encoding a source sequence of pictures comprising the steps of: dividing a source sequence into a set of group of pictures, each group of pictures comprising a first frame, hereafter referred to as I-frame, followed by at least a pair of frames, hereafter referred to as PB-frames; dividing each I-frame and PB-frame into spatially non-overlapping blocks of pixels; encoding the blocks from said I-frame, hereafter referred to as I-blocks, independently from any other frame in the group of pictures; deriving motion vectors and corresponding predictors for the blocks from the temporally second frame of said PB-frame, hereafter referred to as the P-blocks, based on the I-blocks in the previous I-frame or the P-blocks in the previous PB-frame; deriving for each block from the first frame of said PB-frame, hereafter referred to as a B-block, a forward motion vector from said motion vector of the P-block with same location, allowing to obtain for each B-block an associated I-block in the previous I-frame or an associated P-block in the previous PB-frame, hereafter referred to as If-block or Pf-block, respectively; deriving for each B-block of the first frame of said PB-frame, a backward motion vector from said motion vector of the P-block with same location, allowing to obtain for each B-block an associated P-block in the P-frame of said PB-frame, hereafter referred to as the Pb-block; choosing a prediction mode for the encoding of each B-block; predictively encoding the P-blocks of the second frame of said PB-frame based on the I-blocks in the previous I-frame or the P-blocks in the previous PB-frame; predictively encoding the B-blocks following the chosen prediction mode, wherein the choice of the prediction mode for the encoding of each B-block comprises for each B-block in series the steps of: deriving the sum of absolute difference between the B-block and a block with pixels values being the means of the pixels values of the Pb-block and of the Pf-block or If-block, hereafter referred to as SADbidir; deriving the sum of absolute difference between the B-block and the P-block in the second frame of the PB-frame with same location as the B-block, hereafter referred to as SADb; when SADb is greater than SADbidir, making the choice of predictively encoding the B-block based on said P-block with same location as the B-block; when SADb is lower than SADbidir: deriving the difference between said motion vector and said predictor of the P-block in the P-frame of said PB-frame with same location as the B-block; when the obtained difference is lower than a predetermined threshold, making the choice of predictively encoding the B-block based on the P-blocks of the second frame of said PB-frame and the I-blocks or the P-blocks in the previous PB-frame; when the obtained difference is greater than the predetermined threshold, deriving the minimum of the sum of absolute difference for the B-block based on the I-blocks in the previous I-frame or on the P-blocks in the previous PB-frame, and making the choice of predictively encoding the B-block based on the I-blocks or the P-blocks in the previous PB-frame.

2. A system for encoding a sequence of pictures comprising: means for dividing the source sequence into a set of group of pictures, each group of pictures comprising a first frame, hereafter referred to as I-frame, followed by at least a pair of predictively encoded frames, hereafter referred to as PB-frames; means for dividing each I-frame and PB-frame into spatially non-overlapping blocks of pixels; a motion estimator for deriving motion vectors and corresponding predictors for the blocks from the temporally second frame of said PB-frame, hereafter referred to as the P-blocks, based on the I-blocks in the previous I-frame or the P-blocks in the previous PB-frame, for deriving for each block from the first frame of said PB-frame, hereafter referred to as B-block,

a forward motion vector from said motion vector of the P-block with same location, allowing to obtain for each B-block an associated I-block in the previous I-frame or an associated P-block in the previous PB-frame, hereafter referred to as the If-block or Pf-block respectively and for deriving for each B-block a backward motion vector from said motion vector of the P-block with same location, allowing to obtain for each B-block an associated P-block, hereafter referred to as the Pb-block; means for choosing a prediction mode for the encoding of each B-block; means for encoding the blocks from said I-frame, hereafter referred to as the I-blocks, independently from any other frame in the group of pictures, for predictively encoding the P-blocks of the second frame of said PB-frame based on the I-blocks in the previous I-frame or the P-blocks in the previous PB-frame and for predictively encoding the B-blocks of the first frame of said PB-frame, wherein the means for choosing the prediction mode for the encoding of each B-block perform for each B-block in series the steps of: deriving the sum of absolute difference between the B-block and a block with pixels values being the means of the pixels values of the Pb-block and of the Pf-block or If-block, hereafter referred to as SADbidir; deriving the sum of absolute difference between the B-block and the Pb-block, hereafter referred to as SADb; when SADb is greater than SADbidir, making the choice of predictively encoding of the B-block based on the P-block with same location as the B-block; when SADb is lower than SADbidir: deriving the difference between said motion vector and said predictor of the P-block in the P-frame of said PB-frame with same location as the B-block; when the obtained difference is lower than a predetermined threshold, making the choice of predictively encoding of the B-block based on the P-blocks and the I-blocks or the P-blocks in the previous PB-frame; when the obtained difference is greater than the predetermined threshold, deriving the minimum of the sum of absolute difference for the B-block based on the I-blocks in the previous I-frame or on the P-blocks in the previous PB-frame, and making the choice of predictively encoding of the B-block based on the I-blocks or the P-blocks in the previous PB-frame.

3. A system for encoding a sequence of pictures comprising: means for dividing the source sequence into a set of group of pictures, each group of pictures comprising a first frame, hereafter referred to as I-frame, followed by at least a pair of predictively encoded frames, hereafter referred to as PB-frames; means for dividing each I-frame and PB-frame into spatially non-overlapping blocks of pixels a motion estimator for deriving motion vectors and corresponding predictors for the blocks from the temporally second frame of said PB-frame, hereafter referred to as the P-blocks, based on the I-blocks in the previous I-frame or the P-blocks in the previous PB-frame, for deriving for each block from the first frame of said PB-frame, hereafter referred to as B-block, a forward motion vector from said motion vector of the P-block with same location, allowing to obtain for each B-block an associated I-block in the previous I-frame or an associated P-block in the previous PB-frame, hereafter referred to as the If-block or Pf-block respectively and for deriving for each B-block a backward motion vector from said motion vector of the P-block with same location, allowing to obtain for each B-block an associated P-block, hereafter referred to as the Pb-block; means for choosing a prediction mode for the encoding of each B-block; means for encoding the blocks from said I-frame, hereafter referred to as the I-blocks, independently from any other frame in the group of pictures, for predictively encoding the P-blocks of the second frame of said PB-frame based on the I-blocks in the previous I-frame or the P-blocks in the previous PB-frame and for predictively encoding the B-blocks of the first frame of said PB-frame,

9

wherein said encoding system also comprises a computer-readable medium storing a computer program that itself comprises a set of instructions replacing at least some of said means and being executable under the control of a computer or a digital processor in order to carry out, by means of an implementation of the same functions as those fulfilled by the replaced means, the encoding method according to claim 1.

Abstract

In the Improved PB-frames mode, one of the options of the H.263+ Recommendation, a macroblock of a B-frame may be encoded according to a forward, a backward or a bidirectional prediction mode. The invention relates to a method of encoding a sequence of pictures defining a strategy for the choice of a prediction mode among the three possible ones in the encoding of B-macroblock. This strategy is based upon SAD(Sum of Absolute Difference) calculations and motion vectors coherence and allows to use backward prediction when scene cuts occur. The calculations are here performed on original pictures allowing less calculation and reduction in CPU burden. The invention also relates to an encoding system for carrying out said method and including a computer-readable medium storing instructions that allow the implementation of this method.
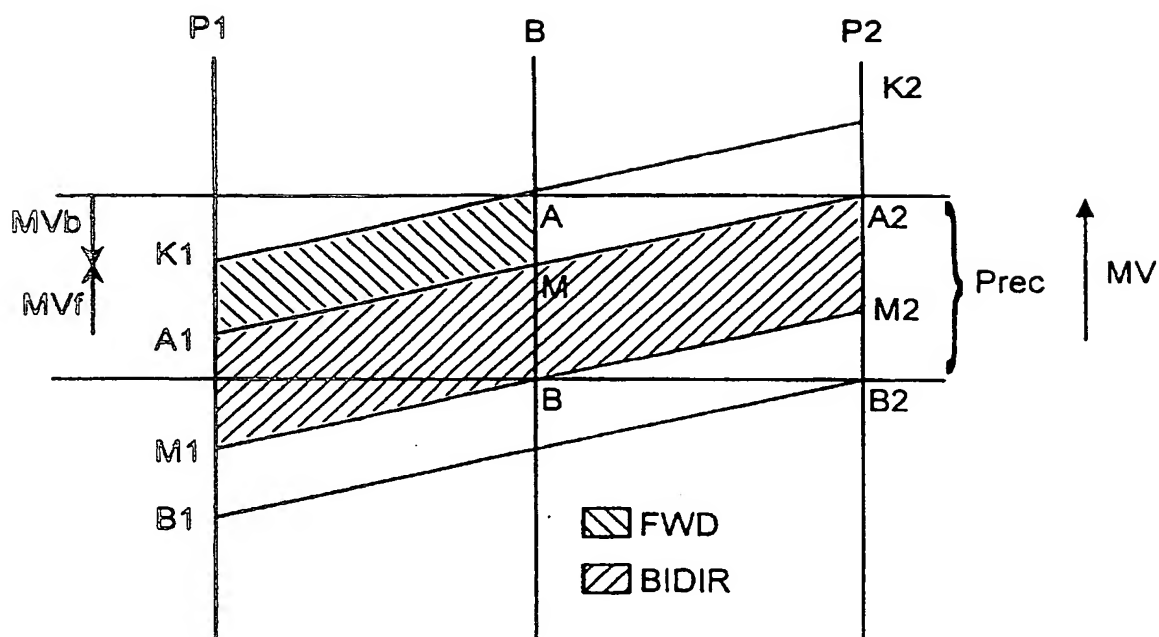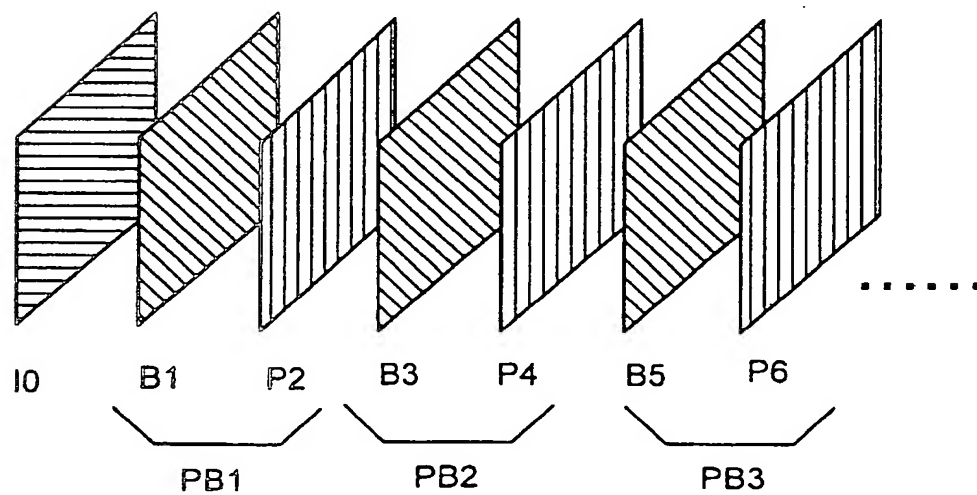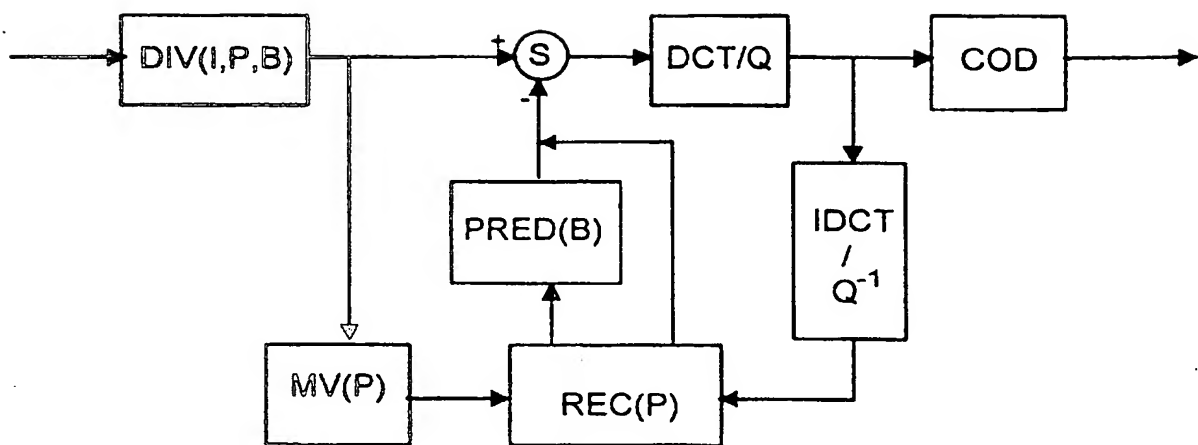
FIG.1

FIG.2

FIG.3

FIG.4

FIG.5